HEIBRIDS
Helmholtz Einstein International
Berlin Research School in Data Science

Admission Round 2019

Project Title	**Deep learning meets structural biology**
Project leads / supervisors	Prof. Dr. Frank Noé (FU Berlin, ECDF) Prof. Dr. Oliver Daumke (MDC, Helmholtz)
Project description	We propose to develop new machine learning methods to significantly improve structural biology approaches to decypher the basic building blocks of life. Over several decades, structural biology, in particular X-ray crystallography and cryo-Electron Microscopy, has brought us probably the most detailed picture of the nanoscale constituents of living systems. However, it is still difficult and involves an enourmous amount of human labour to get high-resolution structures of biomolecules with these methods. Here we want to exploit recent breakthroughs in deep machine learning [5] and build on them to develop specific machine learning methods that help obtaining high-resolution structures of biomolecules. The project will be hosted by the labs of Frank Noé (FU Berlin) who has specialized on statistical and machine learning methods for molecular systems, including deep learning methods [6, 9], recurrent [10] and generative neural nets [11]; as well as Oliver Daumke (MDC) who is an expert for structural biology [1, 7]. The proposed project area actually consists of two specific projects, the choice depends on the candidate and their skills and preferences. In both projects the task is to develop novel machine learning methods that are scalable and achieve statistical optimality in a suitable ensemble. The mathematical formulation of physical and physico-chemical constraints in a deep learning framework (e.g. suitable convolutions to represent certain symmetry groups) is one of the challenges, as well as the implementation of the methods in efficient and publicly available software.
References	[1] K. Faelber, Y. Posor, S. Gao, M. Held, Y. Roske, D. Schulze, V. Haucke, F. Noé, and O. Daumke. Crystal structure of nucleotide-free dynamin. Nature, 477:556–560, 2011. [5] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. Nature, 521:436–444, 2015. [6] A. Mardt, L. Pasquali, H. Wu, and F. Noé. Vampnets: Deep learning of molecular kinetics. Nat. Commun., 9:5, 2018. [7] T. F. Reubold, K. Faelber, N. Plattner, Y. Posor, K. Branz, U. Curth, J. Schlegel, R. Anand, D. Manstein, F. Noé, V. Haucke, O. Daumke, and S. Eschenburg. Crystal structure of the dynamin tetramer. Nature, 525:404–408, 2015. [9] C. Wehmeyer and F. Noé. Time-lagged autoencoders: Deep learning of slow collective variables for molecular kinetics. arXiv:1710.11239, 2017. [10] R. Winter, F. Montanari, F. Noé, and D.-A. Clevert. Learning continuous and data-driven molecular descriptors by translating equivalent chemical representations. ChemRxiv, doi.org/10.26434/chemrxiv.6871628.v1, 2018. [11] H. Wu, A. Mardt, L. Pasquali, and F. Noé. Deep generative markov state models. NIPS (in press). Preprint: arXiv:1805.07601, 2018.